AFRL-OSR-VA-TR-2014-0007

Geometry of Dynamic Large Networks - A Scaling and Renormalization Group Approach

IRAJ SANIEE

LUCENT TECHNOLOGIES INC

12/11/2013
Final Report

DISTRIBUTION A: Distribution approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
04-12-2013

2. REPORT TYPE
Final Performance

3. DATES COVERED *(From - To)*
From 15-09-2011 to 14-09-2013

4. TITLE AND SUBTITLE
TITLE: GEOMETRY OF DYNAMIC LARGE NETWORKS: A SCALING AND RENORMALIZATION GROUP APPROACH
(SUBTITLE: AN ANALYTICAL FRAMEWORK FOR FAST ESTIMATION OF CAPACITY AND PERFORMANCE IN COMMUNICATIVE NETWORKS)

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-11-1-0278

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Iraj Saniee, Mathematics of Networks and Systems, Bell Labs, Alcatel-Lucent
Onuttom Narayan, Department of Physics, UCSC

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
- Bell Labs Research, Alcatel-Lucent, 600 Mountain Avenue, Murray Hill, NJ 07974
-- Dept of Physics, University of California, Santa Cruz, CA 95064

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. Robert Bonneau
AFOSR Complex Networks Program
875 N. Randolph St.
Suite 325
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL

12. DISTRIBUTION / AVAILABILITY STATEMENT
Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The research conducted during the period of performance of this grant concerned investigation of large-scale geometry of prototypical and real-life large graphs, and their analytical/numeric estimation of their core congestion. The research included demonstration of delta-hyperbolicity in a large variety of communication and social networks via scaling of the curvature plots through renormalization [1], detection of graph bottlenecks using a novel techniques based on the Dirichlet eigenvectors of the (normalized) graph Laplacian [2], proof of the fact that remetrization (by factors away from zero and infinity) does not eliminate O(N^2) core congestion in delta-hyperbolic graphs [3], derivation of an analytical expression for the average nodal congestion involving the sum of the inverses of the non-zero eigenvalues of the graph Laplacian when random walk routing replaces geodesic routing [4], and finally determination of the threshold for bootstrap percolation in periodic trees [5] as an upper bounding mechanism for the threshold for more general graphs.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UNL	3	Iraj Saniee
					19b. TELEPHONE NUMBER *(include area code)* 908 582 6410

Final Performance Report

Grant Title: Geometry of Dynamic Large Networks: A Scaling and Renormalization Group Approach

Grant Award Number: FA9550-11-1-0278

Principal Investigators: Iraj Saniee, Mathematics of Networks and Systems, Bell Labs Research, Alcatel-Lucent, Murray Hill, NJ; Onuttom Narayan, Department of Physics, University of California, Santa Cruz, CA

Reporting Period: September 15^{th}, 2011 through September 14^{th}, 2013

Project Accomplishments

The research conducted during the period of performance of this grant concerned investigation of large-scale geometry of prototypical synthetic and real-life large graphs, congestion metrics and analytical estimation of these metrics for such graphs.

In our earlier research (supported by AFOSR grant FA9550-08-1-0064) we had proposed the notion of a Curvature Plot for statistical detection of hyperbolicity, had qualified various aspects of this test, demonstrated $O(N^2)$ scaling of core congestion in delta-hyperbolic networks and provided evidence of hyperbolicity of IP layer of communication networks based on both CP evidence and core congestion. As part of this grant, in "**On the Hyperbolicity of Large-Scale Networks**" [1], we provided detailed analysis of scores of publicly available data sets corresponding to a wide range of large-scale networks, from communication and road networks to various forms of social networks, to explore their hyperbolicity. In smooth geometry, hyperbolicity captures the notion of negative curvature; within the more abstract context of metric spaces, it can be generalized as delta-hyperbolicity. This generalized definition can be applied to graphs, which we explored in this report. We provided strong evidence that communication and social networks exhibit this fundamental property, and through extensive computations we quantified the degree of hyperbolicity of each network in comparison to its diameter. By contrast, and as evidence of the validity of the methodology, we applied the same methods to the road networks to see that they are not hyperbolic according to our tests, which was as expected. Finally, we presented practical computational means for detection of hyperbolicity and showed how the test itself may be scaled to much larger graphs than those we examined via *renormalization group methodology*. Using well-understood mechanisms, we provided evidence through synthetically generated graphs that hyperbolicity is preserved and indeed amplified by renormalization. This allowed us to detect hyperbolicity in large networks efficiently, through much smaller renormalized versions. These observations indicated that delta-hyperbolicity is a common feature of large-scale real-life networks. We concluded that delta-hyperbolicity in conjunction with other local characteristics of networks, such as the degree distribution and clustering coefficients, provide a more complete unifying picture of networks, and helps classify in a parsimonious way what is otherwise a bewildering and complex array of features and characteristics specific to each natural and man-made network.

In "**Spectral analysis of communication networks using Dirichlet eigenvalues**" [2], we shifted attention to detection in bottlenecks in networks, a geometric feature and root cause of congestion in networks. Detection of bottlenecks in graphs is not an easy problem and various methods have been used

historically for this purpose. Among these, the cut ratio, conductance and Cheeger ratio are among the best-known and practiced methods. These are often measured using the Laplacian of the graph. However, graph boundaries often distort these computations. We therefore focused on the spectral gap of the graph Laplacian with *Dirichlet boundary condition* and computed it for the graphs of several communication networks at the IP-layer, which are subgraphs of the much larger global IP-layer network. We showed that the Dirichlet spectral gap of these networks is substantially larger than the standard spectral gap and is likely to remain non-zero in the infinite graph limit. We first proved this result for finite regular trees, and showed that the Dirichlet spectral gap in the infinite tree limit converges to the spectral gap of the infinite tree. We also performed *Dirichlet spectral clustering* on the IP-layer networks and showed that it often yielded cuts near the network core that create genuine single-component clusters. This is much better than traditional spectral clustering where several disjoint fragments near the periphery are liable to be misleadingly classified as a single cluster. Spectral clustering is often used to identify bottlenecks or congestion; since congestion in these networks is known to peak at the core, our results suggest that Dirichlet spectral clustering may be better at finding bona-fide bottlenecks.

In our earlier work, we had shown congestion in an N-node hyperbolic networks scales like O(N^2), as N gets larger, an extreme result since there are no more than N(N-1)/2 undirected flows in a network to begin with. A natural question is whether this result continues to hold in (hyperbolic or small world) networks if we remetrize the graph, that is modify the edge weights. In "**Scaling of Congestion in Small World Networks**" [3], we showed that in a *planar exponentially growing network* consisting of N nodes, congestion scales as O($N^2/log(N)$) independently of how flows may be routed. This is in contrast to the O($N^{3/2}$) scaling of congestion in a *flat polynomially growing network*. We also showed that without the planarity condition, congestion in a small world network could scale as low as O($N^{1+\epsilon}$), for arbitrarily small ϵ. These extreme results demonstrated that the small world property by itself cannot provide guidance on the level of congestion in a network and other characteristics are needed for better resolution. Finally, we investigated scaling of congestion under the geodesic flow, that is, when flows are routed on shortest paths based on a link metric. Here we proved that if the link weights are scaled by arbitrarily small or large multipliers then considerable changes in congestion may occur. However, if we constrain the link-weight multipliers to be bounded away from both zero and infinity, *then variations in congestion due to such remetrization* are negligible.

The above result on the inability of remetrization to reduce O(N^2) scaling of load raises the question of whether non-geodesic routing in some form may change the congestion in significant ways. The extreme opposite of geodesic routing, in some sense, is random routing. In "**Congestion Due to Random Walk Routing**" [4], we derived an analytical expression for the mean load at each node of an arbitrary undirected graph for the uniform multi-commodity flow problem under *random walk routing*. We showed the mean load is linearly dependent on the nodal degree with a common multiplier equal to the sum of the inverses of the non-zero eigenvalue of the graph Laplacian. Even though some aspects of the mean load value, such as linear dependence on the nodal degree, are intuitive and may be derived from the equilibrium distribution of the random walk on the undirected graph, the exact expression for the mean load in terms of the full spectrum of the graph has not been known before. Using the explicit expression for the mean load, we gave *asymptotic estimates for the load* on a variety of graphs whose spectral density are well known. We concluded with numerical computation of the mean load for other well-known graphs without known spectral densities.

Our final result during the performance of this grant concerned investigation of a dynamic process on networks, a generalization of cellular automata to a graph setting. In "**Bootstrap Percolation on Periodic Trees**" [5], we studied bootstrap percolation with the threshold parameter $\vartheta \geq 2$ and the initial

probability p on infinite periodic trees that are defined as follows: each node of a tree has degree selected from a finite predefined set of non-negative integers and starting from any node, all nodes at the same graph distance from it have the same degree. We showed the existence of the critical threshold $p_f(\theta)\in(0,1)$ such that with high probability, (i) if $p > p_f(\vartheta)$ then the periodic tree becomes fully active, while (ii) if $p < p_f(\vartheta)$ then a periodic tree does not become fully active. We also derived a *system of recurrence equations* for the critical threshold $p_f(\vartheta)$ and computed these numerically for a collection of periodic trees and various values of ϑ, thus extending previously known results for regular (homogeneous) trees.

Papers and Publications

[1] Sean Kennedy, Onuttom Narayan, Iraj Saniee, "On the Hyperbolicity of Large-Scale Networks, " http://arxiv.org/abs/1307.0031 (28 Jun 2013), to be submitted.

[2] Alexander Tsiatas, Iraj Saniee, Onuttom Narayan, Matthew Andrews, "Spectral analysis of communication networks using Dirichlet eigenvalues," http://arxiv.org/abs/1102.3722 (7 May 2012) and published in ***Proc. of ACM WWW2013***, Rio de Janiero, May, 2013.

[3] Iraj Saniee, Gabriel Tucci, "Scaling of Congestion in Small World Networks," http://arxiv.org/abs/1201.4291 (20 Jan 2012), submitted.

[4] Onuttom Narayan, Iraj Saniee, Vladimir Marbukh, "Congestion Due to Random Walk Routing," http://arxiv.org/abs/1309.0066 (31 Aug 2013), submitted.

[5] Milan Bradonjic, Iraj Saniee, "Bootstrap Percolation on Periodic Trees," http://arxiv.org/abs/1311.7449 (29 Nov 2013) to be submitted.

Presentations

Onuttom Narayan

- "Curvature of finite graphs and its consequences," NIST-Bell Labs Workshop on Geometry of Large Graphs, Murray Hill, NJ, April 2011.
- "Spectral analysis of communication networks using Dirichlet eigenvalues," WWW2013, Rio de Janeiro, Brazil, May 2013.

Iraj Saniee

- "Large-scale curvature of networks," Advanced Networks Colloquium, A. James Clark School of Engineering, University of Maryland, MD, Oct. 2011.
- "The role of geometry in large-scale network optimization," Plenary Talk at INFORMS 11th Telecom Conference, Boca Raton, FL, March 2012.
- "Large-scale curvature of complex networks and its implications," NIST-Bell Labs Workshop on Geometry of Large Graphs, Gaithersburg, MD, June 2012.
- "Curvature Plots for detection of hyperbolicity in large-scale networks," Santa Fe Institute's Measurements of Complex Networks, Mitre Corporation, McLean, VA, June 2012.
- "Network curvature," NITRD-OSTP Meeting, NSF, Arlington, VA, May 2013.
- "Geometry of large-scale networks and its implications," Santa Fe Workshop on "Structure, Statistical Inference and Dynamics in Networks," Santa Fe Institute, NM, May 2013.